(5) We have seen the mode falls at $x=0.700$, the median at 0.917, and the mean or normal at 1.00.

(6) The probability that a monthly value of rainfall will be greater than the normal is measured by the ratio

$$\frac{\text{area greater than mean}}{\text{whole area}(=A)}=\frac{858.6-481.9}{858.6}=0.44$$

Hence the monthly rainfall will equal or be greater than the normal about 44 months in 100 and of course will equal or be less than the normal 56 months.

(7) For monthly amounts greater than the normal the percentage 1.36 is the probable amount.

(8) *If a monthly rainfall is less than the normal it* will be an even chance that the amount will be greater or less than 66 per cent of the normal, and it was shown under (1) that the most frequent of all monthly amounts was 70 per cent of the normal. Thus it appears that the most frequent monthly amounts and the probable amounts below the average are both about two-thirds the monthly normal.

Such are answers that are easily deduced by interpolations from the mechanical or approximate integration of such scale drawings of frequency distributions as shown in figure 7.

A STATISTICAL COMPARISON OF METEOROLOGICAL DATA WITH DATA OF RANDOM OCCURRENCE.

551.506 : 551.501

By H. W. CLOUGH.

[Weather Bureau, Washington, D. C., Apr. 18, 1921.]

SYNOPSIS.

Daily, monthly and annual means of meteorological data show fluctuations of varying orders of magnitude, which may be regarded as either of a fortuitous character or as presenting more or less systematic characteristics. Certain precise relations which are distinctive of purely fortuitous data are derived by both theoretical and empirical methods. These relations constitute criteria for determining the extent to which meteorological data differ from such fortuitous data.

Monthly and annual means of temperature are nearly Gaussian in their distribution, their deviations being of the nature of accidental errors, but the order of succession of their occurrence is not fortuitous. Rainfall data are more fortuitous in their characteristics than temperature. In a plot of unrelated numbers the two-interval is predominant, while in the case of most meteorological annual means the three-year interval is the most frequent. The variations of mean annual temperatures show systematic characteristics to a greater extent in the Southern Hemisphere and the low latitudes of the Northern Hemisphere than in the higher latitudes of the Northern Hemisphere.

Statistical criteria applied to the variations of the period of the solar spots disclose markedly systematic characteristics.

A period of recurrence of extremes of pressure at Toronto, averaging 32 to 34 days seems to be disclosed by a purely statistical method of treatment of the dates of highest and lowest pressure in each month for a long series of years.

Variability is a dominant characteristic of weather, particularly in temperate latitudes. In the Tropics the day-to-day fluctuations are negligible and the seasonal changes occur with clock-like regularity. The interdiurnal variability of temperature increases with latitude to about the Arctic Circle, then decreases somewhat. A plot of daily mean temperatures exhibits characteristic fluctuations with crests separated by intervals varying irregularly from 3 to 7 days or more. If these daily values be combined into weekly means and plotted there are again shown similar fluctuations but with longer intervals varying from 2 to 5 or 6 weeks. The same data combined into monthly means show, when the residuals are plotted, fluctuations apparently analogous to those of the daily data but with intervals between the successive crests varying from 2 to 6 months or more. Yearly mean temperatures at any locality when plotted show fluctuations which are indistinguishable from a plot of monthly residuals, the intervals being measured in years instead of months.

Thus daily, weekly, monthly, and annual means of meteorological data present fluctuations of varying orders of magnitude. The smaller day-to-day fluctuations are superposed upon the larger weekly fluctuations, the weekly upon the larger monthly, and so on until we arrive at the long secular variations measured by decades or even centuries.

The question arises as to the character of these apparently irregular fluctuations. Are they to be regarded as purely accidental and fortuitous or do they present characteristics which show them to be deviations partaking of a systematic nature and if so, susceptible of prediction? Obviously, if they are of a purely fortuitous nature long-range forecasting is out of the question.

Considerable diversity of opinion regarding this particular aspect of weather changes is gleaned from the literature of the subject. A conservative element regards the monthly, seasonal, and annual variations as due to a complex set of many varying influences whose resultant effect is a series of nearly fortuitous deviations about the normal which can be represented by the well-known Gaussian law of errors. Another element regards the variations as controlled by more or less systematic laws and as being essentially sequences of a quasi-periodic nature. Popular weather lore has for its basis an almost universal belief in the tendency of weather changes to be complementary, in other words for one extreme to be followed by the opposite within a short period.

Obviously it is possible by the employment of statistical criteria to determine the extent to which a given succession of meteorological data conforms to a purely fortuitous selection of similar data, and it will be the purpose of this paper to set forth the characteristics of data which represent purely accidental deviations about a mean and to illustrate by examples of meteorological data how and to what extent the latter differ from data of random occurrence.

CHARACTERISTIC FEATURES OF FORTUITOUS DATA.

There are two classes of data whose deviations present purely fortuitous characteristics: (1) A series of unrelated numbers, illustrated by a random selection of numbers between 0 and 100. In this class of data all values are equally probable. (2) A sample of the component data of a normal frequency distribution, illustrated by the sums of ten digits of random selection. The data of this class are also unrelated, but the various possible values of the variant are of unequal probability.

There are certain precise relations to which these two classes of data rigorously conform and which constitute criteria for testing the conformity of any series of observational data to these requirements. Any deviation from such a conformity indicates some systematic influence operating which results in a frequency curve of either a skew type or a symmetrical but composite type. In either case the curve of best fit by least square methods exhibits excesses in one part of the curve and deficiencies in another part.

Relations between indices of dispersion.—There are various measures of the dispersion or scatter, which, in the case of data showing a distribution of a normal, elemental type, bear mathematically precise relations to each other, which are obviously valid only for a large sample of the component data. One measure of dispersion is the standard deviation σ, which is the square root of the mean of the squares of the deviations, $\sigma=\sqrt{\frac{\Sigma x^2}{n}}$. This value has certain properties which cause it to be extensively employed by statisticians as an index of the dispersoin of the data. The mean and the standard deviation completely determine the form of the Gaussian curve of the best fit to the data.

Another measure of dispersion is the mean deviation or the arithmetic mean of the deviations of the data from the mean, disregarding the signs, $v=\Sigma x/n$.

A third measure of dispersion is the mean variability or mean of the differences between the consecutive values of the variant taken without regard to sign.

A fourth measure of dispersion may be termed the standard variability, computed from the values of variability in a manner similar to that of the standard deviation from the deviations.

A fifth measure of dispersion is the mean of the differences between consecutive maxima and minima, and may be termed the mean range.

These measures of dispersion may be used to determine satisfactorily the characteristics of a series of data as regards the nature of the influences, whether accidental or systematic, operating to control the given data. In a series of data whose deviations are normally distributed these measures of dispersion bear mathematically precise relations to each other, as follows:

(1) *Relation of the mean deviation to the standard deviation, deduced by Cornu.*[1]—Cornu's theorem is "When the departures of a series of numbers satisfy the law of frequency of accidental errors, twice the quotient of the mean of the squares of the departures by the square of the mean departure (mean of the departures made without regard to sign) is equal to π, or in symbols, $\frac{2\sigma^2}{v^2}=3.14159+$.

The standard deviation is thus 1.253 times the mean deviation.

(2) *Relation of the mean deviation to the mean variability.*—Goutereau[2] showed that the mean variability of a series of unrelated numbers, as well as of a series of data whose deviations follow the law of errors, equals the mean deviation multiplied by $\sqrt{2}=1.414$.

Other relations distinctive of unrelated data.—There are certain other relations characteristic of a series of unrelated numbers which form useful criteria for discriminating between purely fortuitous data and data which are subject to some systematic influence and therefore not mutually independent.

(3) *Number of maxima in a series of unrelated numbers.*—Besson[3] showed that the number of crests appearing in a plot of unrelated numbers is one-third of the number of values.

(4) *Frequency of intervals between maxima.*—Of the intervals separating these crests, 40 per cent will be a two-interval, 33⅓ per cent, a three-interval, 17 per cent, a four-interval, 7 per cent a five-interval, and 2 per cent a six-interval. Thus the number of two-intervals is greater than the number of three-intervals in the ratio of 40 to 33.

[1] Cornu, *Annales de l'Observ. de Paris, Memoires.* Tome XIII, p. a220. Paris, 1876.
[2] Ch. Goutereau: Sur le variabilité de la temperature. *Annuaire de la Soci té météorologique de France*, 1906, 54: 122
[3] Besson, Louis. On the comparison of meteorological data with results of chance. (Translation by Edgar W. Woolard.) MO. WEATHER REV., Feb., 1920, 48:89.

The same paper[3] gives the probable number of single rises, double rises, etc., in a series of N numbers. Thus in a series of 100 numbers there are about 20 single rises, and the same number of single falls. It is desirable to know, in addition, the distribution of the single rises or falls as follows: (1) Isolated single rises; (2) single rise followed by a single fall; (3) single rise followed by a single fall followed by a single rise, etc. The total number of single rises and falls in N numbers is $\frac{5N}{12}$. I have deduced the following formula for the number of groups consisting of single rises and falls in a series of N numbers. If n represents the number of single rises or falls in such groups, the general formula is $\frac{5N}{12}\left(\frac{3}{8}\right)^2\left(\frac{5}{8}\right)^{n-1}$. Thus in 1,000 numbers there are approximately 59 isolated single rises and falls, 37 groups with a rise followed by a fall or vice versa, 23 groups with three single rises or falls, 14 groups with four, 9 groups with five.

The writer has derived various other relations which are frequently useful in distinguishing between accidental and systematic deviations.

Select at random say several hundred numbers from a bowl containing perhaps fifty or a hundred of each of the numbers from 0 to 99, inclusive. The mean of these numbers will be 49.5. From the theory of probabilities it can be shown[4] that any two points selected at random on a line of given length are separated by an average interval of one-third the length of the line. Hence the mean variability of the above series of unrelated numbers is 33. The mean deviation is $33/\sqrt{2}=23.3$. In 100 such numbers there are approximately 33 maxima and 33 minima. It is found that the average of the differences between consecutive maxima and minima equals 49.5, or one-half of the possible extreme range. The mean variability is, as stated above, 33. *Hence the mean range is 1½ times the mean variability.* This relation is also valid for the component data of a normal frequency distribution, and is to be classed with the relations deduced by Cornu and Goutereau. The mean range of single rises or falls is approximately 45, or 91 per cent of the mean range of all rises or falls. According to Besson, five-eighths of the intervals between maxima and minima are single rises or falls. This indicates an important characteristic of a series of unrelated data, namely, a relatively large number of wide ranges from a very high to a very low value with no intervening value.

(5) *Distribution of data with reference to mean.*—Another relation characteristic of a series of random data, is the relative frequency of groups of the data above and below the mean. Obviously there are an equal number of values above and below the mean and an equal number of groups, containing 1, 2, 3, 4, or more values, above and below the mean.

If n be the number of values in each group above or below the mean, the probable frequency of such groups in N numbers is $\frac{(N-n-1)}{4(2)^{n-1}}$ or, if N is large, approximately $\frac{N}{4(2)^{n-1}}$. Thus in a series of N unrelated numbers between 0 and 99, there will be $N/4$ separate groups above 49.5 and $N/4$ groups below 49.5, a total of $N/2$ groups. Approximately 50 per cent of the total number of groups above and below the mean will be represented by a single value, 25 per cent by two consecutive values, 12.5 per

[4] *Encyclopedia Brittanica*, 11th ed., Vol. XXII, p. 385.

cent by three values, 6.2 per cent by four values, etc. As illustrating this criterion, the number of consecutive digits 0 to 9, either above or below the mean 4.5, in the fifth place of logarithms of successive integers from 100 to 700 were counted and there were found 4 groups with 14 consecutive digits above or below the mean, 1 group with 18, and 1 group with 19. The theoretical number of groups in a series of 600 unrelated numbers containing 14 consecutive numbers either above or below the mean is 0.02, showing that successive five-place digits in a table of logarithms, while apparently of random occurrence, are actually very far from being unrelated.

(6) *Smoothing formulæ.*—It is instructive to investigate the effect of smoothing formulæ on a series of numbers of random selection. A series of 900 unrelated numbers was smoothed by the formulæ (1) $\frac{a+2b+c}{4}$, (2) $\frac{a+2b+3c+2d+e}{9}$. The number of maxima and minima in the 900 numbers was 598. After smoothing by formula (1) the number was reduced to 384, or 43 per cent, of the 900 numbers, and after smoothing by formula (2) the number was 257, or 29 per cent. Thus the average interval between maxima after smoothing by formula (1) is approximately 4.7, and by formula (2) 7.

The following table gives the approximate frequency of various intervals:—maximum to the next maximum and minimum to minimum—expressed in percentages of the total number of intervals. Column 1 is the interval. Column 2 is the frequency for the unsmoothed data, according to Besson. Columns 3 and 4 give the approximate frequency after smoothing by formulas (1) and (2), respectively.

1	2	3	4
	Per cent.	*Per cent.*	*Per cent.*
2	40	3	2
3	33	30	4
4	17	24	9
5	7	18	15
6	2	13	18
7	0.5	8	16
8	....	4	11
9	....	1	8
10	....	....	6
11	....	....	4
12	....	....	3
13	....	....	2

Thus in the unsmoothed data the two-interval is the most frequent. After smoothing by formula (1) the three-interval is the most frequent, and after smoothing by formula (2) the six-interval is the most frequent.

SYSTEMATIC VS. ACCIDENTAL DEVIATIONS.

It has been stated that meteorological variations show analogies to both accidental and systematic errors of observation. An illustration afforded by target practice may serve to clarify the conception. The center of the target represents the mean and on a calm day the deviations of the shots to the left and right will be equal in number, and symmetrically distributed in conformity with the law of errors. On a windy day the deviations from the center will vary systematically with the direction and velocity of the wind. The shots thus have in addition to purely accidental deviations, a certain systematic deviation which they share in common, so that their deviations from the center are not wholly unrelated or independent of each other. Their deviations in this case will be unrelated only if taken from a new mean whose deviation from the center represents the systematic deviation common to all the shots on the windy day.

In the target practice on a calm day the ratio between the mean variation between the successive shots and their mean deviation will be $\sqrt{2}$. On a windy day, however, assuming the velocity to remain nearly constant from a direction transverse to the line of fire, the deviations from the center will show a systematic increase in magnitude on one side of the mean and a decrease on the other side, while their variability or scatter from each other will be practically unchanged. The ratio, therefore, of the mean variation to the mean deviation will be less than $\sqrt{2}$. If the deviations on all windy days were combined with those on calm days there would result a composite type of frequency curve compounded of two series of deviations each of which have the same variability but very different mean deviations referred to a common mean. This composite curve would be symmetrical but not elemental.

Meteorological data, as will be shown below, are similarly characterized by minor day to day fluctuations, which may be regarded as of the nature of accidental deviations, and larger fluctuations extending over a period of a week or more, analogous to systematic deviations. This tendency for weather to persist in definite types distinguishes any succession of mean daily, monthly, or yearly values of any meteorological element from a series of unrelated numbers.

CITATIONS FROM WRITERS REGARDING THE NATURE OF THE DEVIATIONS OF METEOROLOGICAL DATA.

It was stated above that diversity of opinion exists regarding the nature of the deviations shown by meteorological data. Prominent among those who have written on this subject is Angot,[5] who published in 1900 a discussion of the temperature of France for 50 years. He computed the frequency of monthly departures exceeding $\frac{e}{2}$, e, $2e$, $3e$, and $4e$, where e represents the probable error of the departures. In all cases the actual frequency conformed practically to the theoretical frequency, and, to quote his exact words, "The physical causes that determine one month shall be warm or cold are so many and so complex that the net result is the same as that from purely fortuitous causes." In 1915[6] he made a further investigation of monthly and annual temperatures at Paris from 1851 to 1915 and concluded that "no relation can be made out between the temperature of a season and that of the following season; a warm summer will be succeeded indifferently by a warm winter, or by a cold winter. * * * In conclusion, the variability of monthly, seasonal, and annual temperatures in France follows exactly the same law as if the causes were purely fortuitous and it is not possible to forecast for months, seasons, or years by means of past phenomena."

On the other hand, Goutereau[7] pointed out that the normal ratio between the mean deviation and the mean variability was departed from in the case of certain meteorological data, particularly for a series of successive daily values, indicating the persistence of definite types of weather or systematic deviations from the normal. In such cases the departures may be of the nature of accidental errors but their order of succession is not fortuitous.

---

[5] Angot, A.: Études sur le climat de la France; température, 1re Partie—Stations de comparaison. *Annales, Bur. cent. météorol. de France*, 1897, I. *Mémoires*, Paris, 1899, pp. B93–B170; *ibid.*, 1900, I. *Mémoires*, Paris, 1902, pp. B33–B118.

[6] Angot, Charles Alfred: Sur la variabilité des températures. *Comptes rendus, Acad. Agric. de France*, Paris, déc. 22, 1915, 1:789–792. Transl. by W. G. Reed. MO. WEATHER REV., July, 1916, 44.

[7] Ch. Goutereau: Sur la variabilité de la température. *Annuaire de la Société Météorologique de France*, 1906, 54: 122.

newham

Newham[8] discussed the frequency of "spells" of wet and dry weather at Kew and found that by the law of chance 41 "runs" of 6 rain days should be expected at Kew in 10 years; actually there were 181 "runs" of 6 successive rain days.

Hann (Lehrbuch der Meteorologie, 1915, pp. 629–631) gives results by various investigators showing a tendency for the prevailing type of weather to persist and a decrease in the probability of a change with increasing duration of the type.

Brunt,[9] discussing the monthly residuals of temperature at Greenwich, 1841–1918, says, "It is probable that the greater part of the variations of the monthly means is to be regarded as being of the nature of random variations." He finds that the standard deviation of the monthly residuals corrected for the effect of all the periods found by the Fourier series is very little less than that of the uncorrected residuals and concludes that "the evidence from the standard deviation added to the rather fortuitous manner in which the periods actually formed seem to appear and die away, indicates that investigation of periods in monthly mean temperature is likely to afford very little help in weather forecasting."

The citations from Angot and Brunt fairly represent the views of many meteorologists regarding weather sequences. They either regard the various claims that have been made by investigators, that definite periods exist in weather phenomena, as due to fortuitous combinations that disappear with increasing length of record or quite fail to grasp the significance, from a purely statistical viewpoint, of sequences which can easily be shown to persist over long intervals of time.

STATISTICAL CRITERIA APPLIED TO METEOROLOGICAL DATA.

The problem of weather periodicity may be regarded as presenting two distinct phases. The first question that arises is the extent to which meteorological data exhibit variations analogous to the systematic deviations of physical measurements. Secondly, are these systematic variations periodic or at least quasi-periodic in their nature? In what follows, an attempt will be made to answer the first question by the application of the criteria which have been discussed earlier in this paper and in addition a partial answer to the second question will be developed along purely statistical lines of investigation.

The statistical criteria which have been employed may be enumerated as follows:

(1) Ratio of mean deviation to standard deviation.
(2) Ratio of mean deviation to mean variability.
(3) Relative number of maxima.
(4) Relative frequency of intervals between maxima.
(5) Relative number of groups above and below the mean and relative frequency of total values in each group.
(6) Relative number of maxima and interval-frequency after smoothing.

The numerical values of these various relations for a series of unrelated numbers have been stated above.

Angot confined his investigation to France and a few stations in contiguous countries. I have employed data from the United States, mainly from Schott's discussion of the Smithsonian temperature data.[10]

Angot stated that his conclusions were valid only for France, but it is probable that readers have not always borne in mind this limitation of his results.

TABLE 1.

Stations.	Years of record.	σ	v	u	$\frac{\sigma}{v}$	$\frac{u}{v}$	Maxima and minima.
		° F.	° F.	° F.			Per cent.
Paris	1851–1919	1.11	.91	1.30	1.22	1.42	62
Greenwich	1851–1900	1.15	.90	1.28	1.28	1.42	....
Montpelier	1851–1897	.97	.76	.94	1.19	1.24	....
Marseilles	1851–1897	.90	.77	.85	1.16	1.09	....
Toronto	1841–1908	1.30	1.04	1.29	1.27	1.24	62
Brunswick	1807–1870	....	1.42	1.13	....	.80	56
Salem	1786–1870	....	1.18	1.01	....	.86	62
New Haven	1780–1870	1.25	.99	.97	1.26	.97	56
Philadelphia	1790–1870	....	....	1.05	....	....	60
Baltimore	1817–1904	....	1.15	1.27	....	1.10	59
Cincinnati	1806–1870	1.30	1.01	1.39	1.29	1.38	63
St. Louis	1826–1870	....	1.26	1.55	....	1.23	64
Muscatine	1839–1870	....	1.28	1.64	....	1.28	62
Fort Snelling	1820–1870	1.91	1.59	1.97	1.20	1.24	61
Fort Gibson	1828–1857	....	1.28	1.74	....	1.36	80
Fort Leavenworth	1830–1870	1.78	1.46	2.14	1.24	1.49	68

Table 1 gives for various stations the length of record; the standard deviation, (σ); the mean deviation, (v); the mean variability (u); the ratio, $\frac{\sigma}{v}$, the ratio $\frac{u}{v}$; the number of maxima and minima expressed in percentages of the number of years in the record. The data employed are mean annual temperatures.

(1) The table shows that at most stations the ratio of the standard deviation to the mean deviation averages close to the theoretical ratio, 1.253. This essentially confirms Angot's conclusions and is undoubtedly of universal validity, in the case of monthly and annual means of temperature. He, to be sure, employed a different method for arriving at the same result. Both methods agree in showing that the deviations are distributed about the mean precisely as accidental errors of observation. This implies simply that the mean is the most frequent value and that small deviations are most frequent, and large deviations are relatively infrequent, in conformity with the law of the occurrence of errors.

Marked departures from this ratio signify either (1) a tendency to skewness, (2) a lack of homogeneity in the record, or (3) insufficient length of record. A tendency to skewness is shown by any inequality in the number of positive and negative deviations. As a rule, monthly and annual means of temperature are symmetrical in their distributions.

(2) This, however, by no means tells the whole story regarding the actual sequence of the deviations. When the second criterion, first employed by Goutereau, is applied there is at once apparent a very general departure from the theoretical ratio, $\sqrt{2}$ or 1.414. Paris and Greenwich records have almost exactly this ratio and the Fort Leavenworth record considerably exceeds it. All other stations show smaller ratios. New England stations yield especially low values, the mean deviation exceeding the mean variation. Stations in southern France have markedly lower ratios than at Paris.

A ratio less than the theoretical ratio implies that deviations above and below the mean tend to persist to a greater extent than if they were purely fortuitous. As in the illustration given above, in the case of target practice, a systematic tendency to a persistence of deviations above or below the mean is indicated by a small ratio.

If the data be smoothed by a formula involving at least five consecutive values and deviations be taken from

[8] Newham, E. V.: The persistence of wet and dry weather. *Quart. jour. Roy. meteorol. soc.*, London, July, 1916, 42: 153–162. Abstract in Mo. WEATHER REV., 44: 393.
[9] Brunt, D.: A periodogram analysis of the Greenwich temperature records. *Quart. Jour. Roy. Meteor. Soc.*, London, Oct., 1919, 45: 323.
[10] Atmospheric temperature in the United States. Smithsonian contrib. 277, Washngton, 1876.

newham

these varying means, the ratio between the mean deviations and the mean variations will then closely approximate 1.414.

(3) The third criterion is the relative number of maxima and minima, which in a series of unrelated numbers, is approximately 66 per cent of the number of values.

According to the table, this percentage for annual means of temperature is generally less than 66, ranging usually between 56 and 64. The lowest percentage is in New England, while in the extreme west it exceeds 66 per cent. This excess indicates an excessive preponderance of successive alternations of warm and cold years and is probably confined to the extreme southwest since in Iowa and Minnesota there is the usual deficiency in the percentage.

Meteorological data, as a rule, have less than 66 per cent. For example, the yearly means of pressure at Madras, India, 1841–1919, yield 58 per cent; the yearly means of pressure at Stykkisholm, Iceland, 1846–1918, yield 58 per cent. The monthly residuals of pressure at El Paso, Tex., for 20 years yield 58 per cent.

Rainfall data are, however, more fortuitous in their characteristics than temperature. For example, the yearly temperature at Baltimore, 1817–1904, yields 59 per cent, while the yearly rainfall for the same period yields 69 per cent. The monthly residuals for the same period gave for temperature 60 per cent, for rainfall, 65 per cent.

(4) The fourth criterion classifies the intervals between maxima in percentages of the total number for each 2, 3, 4, 5, etc., interval. For a series of unrelated numbers the normal frequency of each interval is as follows: 2, 40 per cent; 3, 33 per cent; 4, 17 per cent; 5, 7 per cent; 6, 2 per cent; 7, 0.5 per cent.

TABLE 2.

Station.	Years of record.	Intervals. 2	3	4	5	6	7	8	Data.
		P. ct.	P. ct.	P. ct.	P. ct.	P. ct.	P. ct.	P. ct.	
Paris	69	40	29	17	7	5	2	......	Mean annual temperature.
Baltimore	87	30	33	20	14	4	......	......	Do.
New Haven	86	23	39	16	13	8	2	......	Do.
Toronto	68	32	32	30	6	......	......	......	Do.
Interior United States (Schott).	51	33	43	17	7	......	......	......	Do.
Dodge City, Kans.	15	30	46	15	8	2	......	......	Mean monthly temperature.
Madras	79	24	39	24	11	2	......	......	Mean annual pressure.
Stykkisholm	73	21	45	19	10	5	......	......	Do.
St. Louis	48	9	46	14	9	5	13	5	Do.
United States corn yield.	54	28	34	25	12	......	......	......	Mean annual yields.
Unrelated numbers.	......	40	33	17	7	2	0.5	......	

Table 2 gives for selected stations the relative frequencies. As a rule the 3-interval exceeds the 2-interval in meteorological data. At Paris the frequencies for the intervals 2 to 5, inclusive, are practically the theoretical frequencies for unrelated numbers. The 6 and 7 intervals however are two to four times the theoretical values. The preponderance of the 2-interval is an interesting feature, since it is unique in that respect among the stations in the table. At Toronto the number of 4-intervals is nearly equal to the number of 2-intervals. At New Haven the 2-interval frequency is only 58 per cent of the theoretical, while the 5-interval is double and the 6-interval 4 times the theoretical frequency.

The excessive preponderance of the 3 and 4 intervals is very suggestive of a tendency toward a 3 to 4 year period. This is especially marked in the pressure data at Madras and Stykkisholm, where the 4-interval is as frequent as the 2-interval, whereas the theoretical frequency is less than half.

This criterion has been applied to the yearly corn-yields of the United States and the result is given in the table, showing a strongly marked tendency to a three to four year period.

The question arises as to the probability that a longer record would materially change the frequencies given in the table. The values are based on less than 50 years record in many cases. It is found that samples of 50 unrelated numbers vary widely in this respect, some having a preponderance of 3-intervals and a disproportionately large number of 4-intervals. On the other hand where a systematic influence is operating, to cause a preponderance of a 3-interval as compared with the 2-interval, a 50 year record is probably comparable in precision to a sample of 100, or even 150 wholly unrelated numbers.

(5) The fifth criterion is one that brings out more clearly than the preceding ones any tendency toward a persistent deviation from the mean due to some systematic cause. Table 3 gives for various data the relative number of separate groups above and below the mean expressed in per cent of the total number of groups, and the relative frequencies of the number of values in the groups. At Paris the number of groups is 52 per cent of the number of years of record, showing an excess over the theoretical 50 per cent for unrelated data, while the percentage of single years in these groups is 61 per cent as compared with the theoretical 50 per cent. This confirms the previous deduction that the Paris yearly temperatures are very similar to data of random occurrence. The records at New Haven and Toronto, however, yield a much smaller percentage, 32 per cent. Baltimore yearly temperatures for 87 years yield 40 per cent. Schott's consolidated series for the interior United States yields 43 per cent and the Atlantic series yields 39 per cent, both showing a marked departure from a series of unrelated numbers. The monthly residuals of temperature at Toronto for 68 years gives 41 per cent while the monthly residuals of pressure gives 48 per cent, very

TABLE 3.

Station.	Years in record.	Number per group. 1	2	3	4	5	6	7	8	9	10	Number of groups.	Data.
		Per cent.	Per cent.	Per cent.	Per cent.	Per cent.	Per cent.	Per cent.	Per cent.	Per cent.	Per cent.	Per cent.	
Paris	69	61	22	6	3	3	3	3	....	....	....	52	Mean annual temperature.
New Haven	86	41	22	4	7	11	4	4	....	....	4	32	Do.
Toronto	68	36	14	18	18	....	14	....	....	....	....	32	Do.
Baltimore	87	....	....	....	....	....	....	....	....	....	....	40	Do.
United States (interior)	51	41	23	14	13	5	4	....	....	....	....	43	Do.
United States (Atlantic States)	91	43	26	3	11	3	6	6	3	....	....	39	Do.
St. Louis	48	28	28	17	22	5	....	....	....	....	....	39	Mean annual pressure.
Toronto	68	45	24	14	5	5	2	1	2	1	1	41	Mean monthly temperature.
Do.	68	50	25	12	6	2	2	1	1	....	....	48	Mean monthly pressure.
Unrelated numbers	....	50	25	12.5	6.2	3.1	1.6	0.8	0.4	0.2	0.1	50	

nearly the 50 per cent for purely fortuitous data. Bigelow (*Am. Jour. Sci.*, August, 1910) published a table showing temperature departures for the whole United States for each month from 1873 to 1909, inclusive. These monthly departures yield 42 per cent, which is very near the result found for the Toronto monthly residuals. These results show clearly the tendency for prevailing types of warm or cold weather to persist.

The results, however, for pressure as shown by the Toronto 68-year record, shows that pressure variations are more fortuitous in their occurrence than temperature variations. In this respect variations in pressure are much like those of rainfall in showing little indication of systematic deviations in the monthly residuals.

Mielke[11] has compiled departures of yearly temperatures from 1870 to 1910 for 25 districts over the entire globe. The number of stations in each district varies from 4 to 100 or more. The percentage of groups above and below the mean was computed for each district and the results are shown in the following table in which the district percentages are arranged in order of increasing magnitude:

TABLE 4.

	Per cent.		Per cent.
California	30	China and Japan	50
Atlantic States	38	Western middle Europe	50
India	40	Austria	50
Tropical America	43	North Germany and Holland	50
Interior United States	43	Great Britain	51
South Africa	44	Northeast America (eastern Canada)	51
South America	45	South Russia	52
Mediterranean Sea	45	Ural	53
Northwest America	45	Eastern Siberia	53
Australia	46	Northwestern Russia	57
Southwest Siberia	48	Northern Europe	60
Southern United States	49		

There are 12 districts below 50 and 11 districts 50 or above. The districts with percentages below 50 are in the Southern Hemisphere, and the lower latitudes of the Northern Hemisphere, including most of the United States. The low value for the Atlantic States, 38 per cent, is almost identical with the value, 39 per cent, obtained from Schott's consolidated temperatures in the Atlantic States, comprising 91 years from 1780 to 1870. The districts above 50 include northern Europe and Russia. The excessively high values in northern Europe and northwest Russia illustrate the extreme variability of weather in high latitudes. According to Besson the number of single rises and falls in a series of 100 unrelated numbers is approximately 41 while in the district of northern Europe the percentage is 52, showing a tendency to a two-year period.

It is obvious that a marked deviation either above or below 50 per cent is indicative of systematic tendency in the variations. These results are interesting in showing how different are the characteristics of meteorological variations in different regions, and how unsafe it is to draw general conclusions from investigations covering a restricted area.

(6) The sixth and most significant criterion relates to the effect of mechanically smoothing the data by certain formulæ. The following table gives the relative frequency of the various intervals after smoothing by the formula $\frac{a+2b+c}{4}$ for the data (1) Schott's United States interior temperature; (2) Toronto mean annual temperature; (3) Paris mean annual temperature; (4) a series of unrelated numbers.

TABLE 5.

Interval.	(1)	(2)	(3)	(4)
2	....	5	4	3
3	7	18	23	30
4	13	18	27	24
5	27	18	23	18
6	33	9	8	13
7	13	0	8	8
8	7	22	4	4
9	....	5	4	1
10	....	5	....	....
	35	35	40.6	43

The numbers at the foot of the table are the relative number of maxima and minima in the smoothed data expressed as a percentage of the number of years in the record. This table shows clearly the systematic tendency to persistence of the same type of weather even at Paris where by Angot's and other criteria the data are nearly indistinguishable from data of random occurrence. Thus at Paris the 4-interval is the maximum and the 3 and 5-intervals are equal while the theoretical frequency gives the 3-interval as the maximum and the 5-interval is little more than half the 3-interval. The marked departure of the frequencies for Schott's temperatures, interior of United States, from the theoretical frequencies is a particularly striking feature, in view of the fact that the percentage of maxima and minima in the unsmoothed data is 67.

The interpretation of these results evidently is that the amplitude of the minor year-to-year fluctuations is relatively small compared with that of unrelated numbers, where, as shown above, the amplitude of the single rises or falls is but little less than the average amplitude of all rises and falls—45 as compared with 50. By smoothing, the small fluctuations of temperature disappear, leaving a relatively small number of the large fluctuations.

This criterion is of greater value than any of those previously mentioned for disclosing systematic tendencies in a series of observational data. The Schott temperatures yield by the application of the third criterion—that of Besson—a result which is identical with that of a series of unrelated numbers. The same is true of the Paris temperatures. By smoothing, the essentially systematic character of the data is clearly brought out.

STATISTICAL CRITERIA APPLIED TO SOLAR DATA.

The criteria above mentioned are applicable in all cases where it is desirable to determine whether systematic deviations are present in any given set of data. As an illustration of the value of such criteria the sunspot epochs of Wolfer have been examined and the results are given below. The variations of the 11-year period have long been recognized. Newcomb[13] discussed the variability of the period and concluded from a mathematical analysis of the data that the deviations in the length of the period from a normal period were of an accidental nature. Some extracts from his paper follow:

> In discussing periodic phenomena in which the times of recurrence of a given phase are subject to irregularities, two hypotheses may be made. One is that underlying the periodic phenomena there is a primary cause going through a perfectly uniform period, but that on the action of this cause are superseded irregular actions which may delay or accelerate the occurrence of a phase without affecting the primary cause. When this is the case we shall have a series of perfectly equidistant normal epochs for the recurrence of the same phase, and the observed deviations from these epochs will be in the nature

[11] Mielke, Johannes: Die Temperaturschwankungen, 1870-1910. *Archiv. des Deutschen Seewarte*, XXXVI, 1913. Hamburg, 1913.

[13] Newcomb: The period of the solar spots. *Astrophysical Jour.*, Jan., 1901, 13:1.

of separate and independent accidental errors. If p be the true value of the normal period, then at the end of n periods, however great n may be, the time of occurrence of the phase will differ from np by a small quantity $\pm e$ indicating the irregularity in the general mean. This value of e will be the same no matter how great n may be.

The other hypothesis is that while there is still a certain normal mean period, this period is nevertheless subject to change in such a way that if a phase is once accelerated the advance thus produced will go on indefinitely into all subsequent phases.

By a least square solution he deduced the mean period to be 11.13 years and concluded that the deviations of the epochs from the mean epochs based on this length of the period were of an accidental nature and that the first hypothesis was the correct one.

In order, however, to justify the adoption of this hypothesis he found it nesessary to regard some of the observed phases as more or less erroneous, due to the imperfections in the record. He writes:

I think that these perturbations of the period about 1790 are to be regarded as errors rising from the imperfection of the record. * * * The fact appears to be that while modern observations show that the maximum follows the minimum by less than 5 years and between 6 and 7 years are required to again fall to the minimum, the older observations seem to place the two epochs nearly equidistant. I regard this only as resulting from the accidental errors of the observations, as we can scarcely suppose a change in the law of variation to have occurred. * * * The contrast between the sudden deviations in the residuals of the doubtful period and the small ones of the recent well-observed epochs, make it almost certain that the errors between 1770 and 1800 are due to imperfections of the record.

There are, as Newcomb pointed out, two classes of errors. One is that of the observations themselves; the other, the irregularities of the actual phase. He regarded both classes of errors as of an accidental nature.

It is obvious, therefore, that Newcomb found it necessary to cast doubt upon the accuracy of the epochs in order to establish his theory of accidental variations. The actual deviations were somewhat greater than the theory allows.

Clough (*Astrophysical Journal*, vol. 22, No. 1, p. 62), discussed this point and concluded from several converging lines of evidence, including the testimony of magnetic and auroral data, that the epochs of Wolfer were substantially reliable. Wolfer himself has found it necessary to reiterate that the accuracy of the epochs in the latter part of the 18th century was greater than some students seemed inclined to allow. He considers the variations in the length of the period to have apparently a periodic character.

The problem is thus one of considerable historic interest. Fortunately it lends itself readily to a purely statistical treatment and furnishes a peculiarly apt illustration of the value of the criteria which have been employed above on meteorological data.

The mean epochs for the maxima and minima have been formed by extending backward Newcomb's mean epochs. The average deviation of the observed from the mean epochs are for the maxima, 1.40 years and for the minima, 1.13 years, showing a much greater precision for the determination of the epochs of minimum. The mean variability of the deviations of the maximum phases is 1.56 years, and of the minimum phases, 1.27 years. The ratio between the mean deviation and the mean variability is 1.12 for both maximum and minimum deviations, showing a marked departure from the ratio 1.414, which would obtain if the deviations were of a fortuitous character.

Applying the fifth criterion there is found for each of the series of deviations of the maxima and minima that the number of separate + and − groups is 36 per cent of the number of values. This is markedly less than the theoretical 50 per cent for purely accidental deviations.

The sixth criterion, based on smoothed values, has been employed on the sunspot intervals and the results are a striking demonstration of the value of this criterion for distinguishing between accidental and systematic deviations.

In my paper, previously referred to, the successive intervals maximum to maximum and minimum to minimum were combined and smoothed by the formula $\frac{a+b+c}{3}$. The resulting values were plotted on Chart I of that paper, and showed fluctuations of a more or less regular character. In order to apply the criterion given above in connection with the formula $\frac{a+2b+c}{4}$ it will be merely necessary to smooth the successive 11 year intervals by the formula $\frac{a+b}{2}$. Regarding the values a and b or the interval from minimum to maximum and maximum to minimum, respectively, as the elementary data to be examined, it is obvious that the smoothing of the 11-year intervals, $a+b$, $b+a'$, $a'+b'$, etc., by the above formula gives smoothed values of a and b by the formula $\frac{a+2b+c}{2}$. It might seem that, since the value of b is systematically greater than of a in the ratio of 1.2 to 1.0, taking the average from 1610 to 1900, or according to Newcomb, 1.4 to 1.0, there would be an alternation of higher and lower values in the smoothed numbers. However, the actual deviations of the occurrence of the phases are of an order of magnitude that considerably exceeds the relatively smaller variations due to the systematic differences between a and b. There are in the smoothed values 19 maxima and minima, or 17, if a slight fluctuation showing an increase from 10.7 to 10.75 be disregarded, since the amplitude of this fluctuation is less than 1 per cent of the larger fluctuations. The total number of values is 54. $\frac{17}{54}$ is 32 per cent which is to be compared with 43 per cent, as stated above, if the data were purely fortuitous.

Classifying the intervals according to magnitude there are obtained the following frequencies expressed in percentages of the total number of intervals. Column (1) is the magnitude of the interval. Column (2) gives the theoretical frequency for fortuitous numbers.

(1)	(2)	(3)
2	3	..
3	30	..
4	24	27
5	18	13
6	13	20
7	8	20
8	4	13
9	1	0
10	..	7

Column (3) gives the relative frequency for the smoothed sunspot intervals. The mean interval length is 4.7 for purely fortuitous numbers. The average interval for the sunspot numbers is $\frac{54\times}{17}=6.4$, or 35.2 years, since the unit interval is approximately 5.5 years.

The application, therefore, of these simple criteria shows conclusively that the deviations of the epochs of maxima and minima, instead of being accidental as Newcomb concluded, are systematic to a marked degree, to a greater extent indeed than in the case of any series of terrestrial meteorological data yet examined.

QUASI-PERIODIC NATURE OF THE SYSTEMATIC VARIATIONS DISCLOSED BY STATISTICAL METHODS.

In what precedes, systematic tendencies to a persistence of weather types have been shown to characterize mean values of meteorological elements, particularly temperature. The question arises as to the periodic or quasi-periodic nature of these systematic variations. It is possible by purely statistical methods to show that there is a marked preponderance of certain intervals of recurrence of like phases in these variations. For example it has been shown that the three-interval is the most frequent interval in most series of meteorological data and some records, for example the mean annual pressures at Madras and Stykkisholm show such a preponderance of the three-interval that a three-year period may be said to be the most prominent one in annual means. I have not employed to any extent in this discussion data outside of the United States, but it is well known that in the Tropics and in middle latitudes of the Southern Hemisphere, the three-year period is clearly obvious from mere inspection of the plotted unsmoothed data, and it is scarcely necessary to apply any of the criteria which have been employed on data in high latitudes. With increasing distance from the Tropics, especially in the Northern Hemisphere, the fluctuations become more irregular and finally become nearly indistinguishable from those of unrelated data. The necessity for smoothing the data then arises in order to bring out any tendency to systematic variation. The fact that in certain regions, which correspond closely to the belt of maximum storm frequency, the variations are of a nearly fortuitous character, and there is a marked seasonal inequality in their systematic characteristics, corresponding to the seasonal shifting in latitude of the storm belt, points to the obvious conclusion that investigations of these systematic variations, which renders possible long-range forecasting, should first be restricted to regions in low latitudes and in longitudes relatively free from cyclonic action. Thus in the United States, the Southern States and the region west of the Rocky Mountains should exhibit more regularity in the long period fluctuations than other regions.

It should furthermore be borne in mind that monthly residuals of rainfall are more nearly fortuitous in their occurrence than temperature data, so that it would be rash to draw conclusions from an examination of rainfall data alone. The haphazard character of rainfall, particularly in summer, is well known.

It is obvious that the segregation of long period fluctuations from those of shorter duration, which is facilitated by a smoothing process, is the preliminary step in the investigation of meteorological data. In many cases simple inspection of the smoothed data, bearing in mind the criteria based on smoothing formulæ, discloses a more or less regular periodicity. Thus the curves illustrating the 7-year cycle (Cf. Maurer, *Archives des Sci. Phys. Nat.* Geneva, May, 1918; and Clough, Mo. WEATHER REV., Oct., 1920, **48:** 593–597) speak for themselves when it is remembered that in the case of fortuitous data, similarly smoothed, the most frequent interval is the three-interval, and the average interval is 4.7. It is difficult to understand how any other possible treatment of the data can strengthen the conclusions which necessarily follow from simple inspection of smoothed data in which periodic recurrences are so clearly evident as in the case of the 3-year and 7-year intervals.

The 35-year Brückner cycle is another instance of the determination of a periodic recurrence by inspection of curves of smoothed data. Brückner employed a slight modification of the ordinary smoothing process which leads to similar results. For exhibiting a 35-year cycle, 5-year means sufficiently smooth out the minor fluctuations to disclose the larger periodicity.

The monthly periodicity.—The existence of a weather periodicity apparently coinciding with the synodic rotation of the sun or the synodic revolution of the moon has been affirmed by many and there is an extensive literature dealing with the subject. One may be tempted to think it incredible, in view of the magnitude of research which has been expended on this problem, that there should be no real foundation for such assumptions. It is not necessary, however, to assume that either the solar or lunar periods are directly related to the supposed period, whose length may be sufficiently near either of these periods to lead to the plausible assumption that a real solar or lunar influence was in operation.

Passing over consideration of the vast literature of the subject, extracts will be given from a paper by Koeppen [13] summarizing an exhaustive investigation which he made in examination of the claims of lunar periods in the weather. After pointing out the fallacies commonly met with in such claims, he writes:

> Nevertheless, the application of correct methods has brought out several points wherein there are signs of a lunar influence, and these must be further investigated. On the one hand these signs indicate an atmospheric tidal movement, very slight, to be sure, and of infinitesimal effect upon weather and wind, as are the daily barometric variations in any case. On the other hand they point to more or less considerable fluctuations of about one month's duration; the regularity of these swings leaves it an open question whether they belong with one of the periods of the lunar revolution or of the sun's rotation, for these have similar durations.

He investigated pressure variations in Europe between 1755 and 1912 and concludes that while there is some systematic tendency for a recurrence of similar conditions about a month apart, there is no evidence of any regularity or persistence of a definite relation to the lunar synodic period throughout the entire series.

The concurrence of much evidence indicating a period of about one month led the author to develop a statistical method by which it seems possible to prove or disprove the existence of a period of such length. This is based on the fact that if there is selected alternately at random a number between 1 and 30 and a number between 31 and 60, the mean *difference* between these two numbers will be 30 and will be distributed symmetrically about this mean. If, therefore, we have a record of the dates of highest and lowest pressure for each month for, say, 50 years, at any locality and take the intervals between the dates of highest or of lowest pressure for two consecutive months, a month being approximately 30 days in length, these *intervals*, if the dates are of purely fortuitous occurrence, will have a maximum frequency of 30 days and the intervals above and below 30 days will decrease in frequency to a minimum for intervals of 0 and 60 days.

There were available the dates of highest and lowest pressure at Toronto for each month from 1840 to 1915. The intervals were tabulated separately for the years 1840–1879 and 1880–1915, also for the months April to September and October to March.

[13] *Meteorologische Zeitschrift*, 32: 180-185, April, 1915. (Translated by C. Abbe, jr., in Mo. WEATHER REV., April, 1915.)

The following table gives for the warmer half of the year the relative frequencies of the intervals for each 10 days between 10 and 50 days:

Table 6.

	10-19	20-29	31-40	41-50
	Per cent.	*Per cent.*	*Per cent.*	*Per cent.*
1840-1879	16	26	34	23
1880-1915	17	26	34	23

The two halves of the period of 75 years show practically identical distribution. It is clearly obvious that a markedly unsymmetrical distribution has persisted throughout the entire period in summer, 43 per cent of the total number of intervals between 10 and 50 being below 30 and 57 per cent above 30. The mode for the entire period is approximately 34.

The following table gives the relative frequencies for the winter months:

Table 7.

	10-19	20-29	31-40	41-50
1840-1879	19	29	32	20
1880-1915	16	34	33	16

Thus for the entire period the distribution is practically symmetrical, with the mode slightly above 30.

The following table gives the interval frequency for the dates of maxima and minima separately for the two halves of the year:

Table 8.

	10-19	20-29	31-40	41-50	10-29	31-50
MAXIMA.						
Winter	19	28	33	20	47	53
Summer	17	27	32	24	44	56
MINIMA.						
Winter	16	35	32	17	51	49
Summer	16	24	37	22	40	59

Examination of these figures shows that in the warmer half of the year the distribution of the intervals for both maxima and minima is of marked asymmetry. In winter, however, the intervals based on the dates of minima are of practically symmetrical distribution, while the maxima yield a distribution with a slight tendency to asymmetry, not, however, so pronounced as in summer.

The difference between the results for winter and summer may be plausibly accounted for by the well-known tendency for HIGHS and LOWS to be of more intense development and rapid movement in winter. The extreme pressures are confined to a much smaller area when the systems are intensely developed and consequently a more nearly fortuitous occurrence of the dates of extreme pressures at any one locality would result. Low pressure areas, being more variable in their departures from the normal than high areas and with the extreme reading more localized, there would naturally result a greater tendency to fortuity in the dates of occurrence of lowest pressure.

Suppose, for example, in addition to the Toronto data we had similar data for Rochester. We should expect to find, as in fact we actually do, more agreement between the dates at the two places in summer than in winter and in winter there would be closer agreement between the dates of maxima than of minima.

Thus if there is a tendency for systematic recurrence at intervals somewhat greater than 30 days, as seems to be indicated by the results for the warmer season, this tendency would be modified or even entirely obliterated by the greater tendency to fortuitous occurrence in winter, particularly for extremes of low pressure.

A further compilation was made of the intervals between the dates of maximum pressure in each month and that of the second month following. The most frequent interval, if the dates were of purely fortuitous occurrence would be 60. Actually the most frequent interval was around 65, which is a double 32 to 33-day interval.

It should be understood that the results by this method are not to be interpreted as indicating the probable length of the monthly periodicity with any degree of accuracy. All that may be reasonably deduced from the facts here presented is that there is a systematic tendency for the recurrence of periods of high pressure, particularly in summer, at intervals somewhat greater than 30 days. The tendency for a purely fortuitous occurrence of the dates in winter, particularly so for the dates of minima, is what we should, *a priori*, expect. This being the case, a marked departure from a symmetrical distribution, which occurs in summer and has persisted for 75 years as shown by the close agreement of the results for the two halves of the period, can not be explained other than as a result of a systematic tendency for the dates of extremes of pressure to depart from a purely fortuitous occurrence. The question as to the actual average length of the period and to what extent it may vary in length from time to time is left unanswered. Other evidence, however, indicates that this periodicity may have a variable length over a range of a week or more and hence investigators who have observed recurrences which they regarded as of solar or lunar origin, may have been misled by the apparent coincidence of a minimum length of the monthly periodicity with solar or lunar periods. When the period resumed its normal length the apparent coincidence disappeared. Thus Koeppen's results which require for their explanation the hypothesis of a systematic tendency to a monthly periodicity are plausibly explained by variations in the length of the period.

551.501

THE MEAN VARIABILITY AS A STATISTICAL COEFFICIENT.

The difficulty of applying the ordinary Theory of Errors to meteorological computations, on account of the peculiar nature of the meteorological variables as contrasted with that of the mathematical variables,[1] has often been recognized.[2] If the arithmetic mean of a series of values is to be the value most worthy of confidence, and is to have any significance and correspond to something physical, then the individual values from which it is computed must be distributed about it according to the Law of Gauss—the deviations from the mean must obey the laws of fortuitous errors.[3]

There are two equivalent tests which are ordinarily applied in order to determine whether or not the individual deviations from the mean are due to fortuitous

[1] L. Besson: On the comparison of meteorological data with results of chance. Mo. WEATHER REV., Feb., 1920, 48: 89.

[2] V. H. Ryd: On computation of meteorological observations, *Danske Meteorologiske Institut*, 1917.

[3] Angot: *Annales du Bur. Cent. Mét.*, 1895 and 1900; and *Annuaire de la Soc. Mét.*, 51, 1903.